Best Ways to Use Billions of Devices on a Chip – Error Predictive, Defect Tolerant and Error Recovery Designs

Kazutoshi Kobayashi[1] Hidetoshi Onodera[1,2]
1. Graduate School of Informatics Kyoto University, 2. JST, CREST
e-mail: {kobayasi, onodera}@vlsi.kuee.kyoto-u.ac.jp

Abstract— Error rates on an LSI are increasing according to the Moore's law. Now is the time to start incorporating error-tolerant design methodologies. This paper introduces sources of failures in semiconductor devices, levels of dependability according to applications of devices and some circuit-level techniques to detect or recover faults after shipping.

I. Introduction

Billions of transistors on a single die change our daily life more convenient year by year. Everything can or will be replaced by semiconductor devices. For example, CRT displays, wireline telephones and personal cassette-tape players (Walkman) are replaced with LCD/TFT displays, cellphones and silicon-audio players (iPod), respectively. Replacing with semiconductor devices makes everything portable. We can not carry CRT displays but cellphones with TFT displays, built-in portable TVs and silicon audio players.

As the growth of consumer electronic devices, our daily life highly relies on portable (wireless, battery-powered) and unportable (wireline, ac-powered) devices and systems. All kinds of informations are concentrated on cellphones (portable devices) such as phone numbers, e-mail addresses, electric commerce, e-cash and so on. On the other hands, so many unportable systems take control of our social infrastructures like e-commerce or on-line reservation services. Breakdown of portable devices may just lose personal information, which damage is small. But the systems used for the social infrastructure must avoid sudden breakdown. In 20 years ago, breakdown of ATMs (Automated Teller Machines) will be covered by human tellers. But now it is almost impossible to operate banking systems without ATMs and computer networks. Recently, a Japanese airline, ANA has major troubles on their reservation system caused by failures on a network switch. Because of the trouble, ANA canceled 130 flights and over 300 flights were delayed. About 70,000 people were affected. Fault-tolerant design methodologies must be incorporated to devices for infrastructures and personal use. But the amount of dependability are different from application to application.

II. Failures on Semiconductor Devices

According to the Moore's law, number of transistors on a die are tripled every two or three years, which also increases the probability of permanent or temporal failures of devices. The possibility of failures on Gb DRAM in 2007 is 10^3 higher than that on Mb DRAM in 1985. We must design semiconductor devices with capabilities of error prediction, error tolerance or error correction on VLSIs with billions of Trs.

Failures on semiconductor devices are categorized as follows.

Temporal Failure Functionality of circuits are temporarily failed. Soft errors caused by alpha particles is one possible source.

Permanent Failure Functionality of circuits are permanently failed. Electromigration is one possible source.

Aging Degradation Performance of devices are deteriorating gradually with age. NBTI (Negative Bias Temperature Instability) and HCI (Hot Carrier Injection) are possible sources. Unlike the permanent failure, circuits will operate under relaxed conditions (Higher Vdd or lower clock frequency).

TABLE I
Levels of Dependability According to Applications.

Applications	Cellphone	Car	Infrastructure
Cost	Low	Middle	Huge
Tempo. Fail.	Acceptable	Unacceptable	
Parm. Fail.	Best Avoided	Unacceptable	
Aging Degra.	Acceptable	Unacceptable	Acceptable

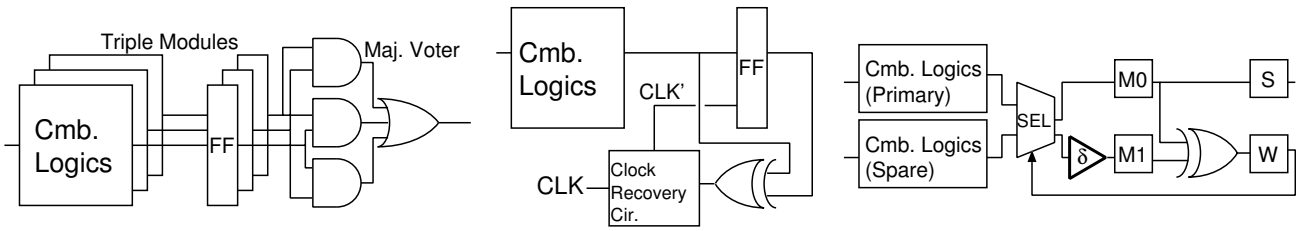
III. Levels of Dependability According to Applications

VLSIs are currently used for various applications. Table I shows three typical applications, cellphone, car and infrastructure which demand different levels of dependability. Cellphones allow temporal errors that can be fixed by cycling power. It is better to avoid permanent errors but replacing after permanent breakdown may be acceptable for cellphones. But devices for cars and infrastructures must avoid both temporal and permanent failures for safety. Devices for cellphones and infrastructure accepts some amount of aging degradation. No one uses cellphones for 10 years. Devices for infrastructure are placed in stable ambient conditions, in which aging degradation can be precisely predicted at a design time. But devices for cars are exposed to totally different conditions. Some car has lifetime less than 5 years but another keeps on running over 10 years. Ambient conditions change drastically for car devices. If the worst conditions are assumed, timing margins at a degign time to compensate the aging degradation will become huge. It may be better to compensate the degradation after shipping.

978-1-4244-1922-7/08/$25.00 ©2008 IEEE

TABLE II
THREE CIRCUIT-LEVEL TECHNIQUES AGAINST FAILURES

	Conv. FF	TMR [1]	CFR-FF [2]	DP-FF [3]
Area Ratio	1	3	$+1-2\%$	2.5
Delay Overhead	-	10%	2%	3%
Error Rate	α	α^2	0.5α	-
Tolerance	-	Temporal Error		Permanent Error



(a) Triple Modular Redundancy (b) CFR-FF (c) DP-FF

Fig. 1. Circuit Schematics of TMR, CFR-FF and DP-FF.

IV. CIRCUIT-LEVEL TECHNIQUES AGAINST FAILURES.

Recently, so many circuit-level techniques are introduced to recover or predict failures on VLSI with some amounts of space or time redundancy. Table II lists several error-tolerant design techniques and conventional circuits without redundancy (referred as Conv. FF). Triple Modular Redundancy (TMR) [1] prepares triple equivalent circuits and/or FFs as shown in Fig.1(a). The voter circuit chooses a correct result by a majority vote among three results. Error Rate of TMR is drastically reduced to α^2 compared with α of the circuit with no redundancy, since a TMR circuit keep on running until two modules fail at the same time. But its area penalty are large.

Compact Fault Recovering (CFR) Flip-Flop [2] as in Fig.1(b) detects discrepancy between input and output of a FF. When an error is detected, an additional clock is given to the FF using the clock recovery circuit within the same cycle to recover errors. Its area penalty is very small but the error rate is reduced only by half of the conventional FF. Defect-Prediction FF [3] is designed to compensate permanent errors mainly caused by electromigrations (called as latent defect-related failures in [3]). Two combinational logics are prepared. The primary cmb. logics is working after shipping while the spare cmb. logics is turned off. If electromigrations gradually worsen the delay of the primary circuit, two master latches (M0 and M1) holds different values. Then the XOR gate becomes high and the spare circuit starts to work instead of the primary one. It will prolong the MTBF at the cost of $\times 2.5$ area penalty. In order to recover temporal errors, a method using space redundancy such as TMR is the best solution without considering the area penalty. But the error rate of α^2 is unnecessarily small. On the other hand, a method using time redundancy such as CFR FF achieves small area, but the order of the error rate is almost same as the conventional FFs. Circuit techniques are required to achieve less area penalty and modest error rate. Dual modular redundancy (DMR) reduces area penalty to twice but cannot recover errors without latency overhead for rollback.

Aging degradation are currently considered at the design time. Extra timing margins including the delay increase by aging are added to the worst case delays. Currently no correction or compensation techniques are applied at the post-silicon phase. But aging degradation can be compensated after shipping. The degradation is gradually increased with time. Prediction may be done when devices are turned on, that can avoid delay penalty during working. The dual or triple modular techniques are ineffective for aging degradation since working modules degrade their performance.

V. CONCLUSION

Levels of dependability are different from application to application. Devices for infrastructures must avoid temporal failures, while those for cellphones may accept sudden breakdown by temporal failures. We must find the best way to mitigate failure rates considering the penalties of area, delay, power and so on.

REFERENCES

[1] L. Anghel, D. Alexandrescu, and M. Nicolaidis. Evaluation of a Soft Error Tolerance Technique Based on Time and/or Space Redundancy. *sbcci*, 00:237, 2000.

[2] Shin'ichi Yasuda and Shinobu Fujita. Compact Fault Recovering Flip-Flop with Adjusting Clock Timing Triggered by Error Detection. *Custom Integrated Circuit Conference*, pages 721–724, 2007.

[3] T. Nakura, K. Nose, and M. Mizuno. Fine-Grain Redundant Logic Using Defect- Prediction Flip-Flops. *ISSCC*, pages 402–403, 2007.